

令和 4 年 6 月 6 日現在

機関番号：34310

研究種目：若手研究(B)

研究期間：2017～2021

課題番号：17K12749

研究課題名（和文）相転移描像による圧縮センシングの限界解明とその打破

研究課題名（英文）Breakthrough of limits in compressed sensing by the picture of phase transition

研究代表者

中西 義典（Nakanishi, Yoshinori）

同志社大学・文化情報学部・准教授

研究者番号：00767296

交付決定額（研究期間全体）：（直接経費） 3,200,000円

研究成果の概要（和文）：スパースモデリングによる圧縮センシングを未知の対象に適用するという現実的な状況で、その成否を判断する手法を交差検証の考え方に基ついて開発した。統計力学のレプリカ法を用いて解析すると、データ量が増えて圧縮センシングが失敗から成功へと転移するところで交差検証誤差が冪的に振舞うことが分かった。交差検証の訓練データと検証データの比率を変えた時に見られる交差検証誤差の振る舞いを調べることにより圧縮センシングの成否を判断する手法を提案し、その性能を数値実験により評価した。

研究成果の学術的意義や社会的意義

本研究の成果として、圧縮センシングの限界が明らかになったことにより、不十分なデータによる誤った解析結果を鵜呑みにすることなく、またデータが十分であることを認識できずに過剰なデータ取得に陥るようなこともなくなると考えられる。本研究で開発した手法は交差検証に基ついており、スパースモデリングに限らずあらゆる情報学的手法と組み合わせることができる。圧縮センシングの他にも情報学的手法を用いた実験・計測の効率化は試みられており、そのような場合にも発展する可能性がある。

研究成果の概要（英文）：This research developed a cross-validation based method for judging the success or failure of compressed sensing, considering practical situations where compressed sensing with sparse modeling is applied to unknown objects. Analysis with the replica method from statistical mechanics revealed a power-law behavior of cross-validation error at the transition point in the amount of available data from the failure to the success of compressed sensing. This research proposed judging the success or failure in compressed sensing based on the dependence of cross-validation error on the amount ratio between training and validation datasets in cross validation and evaluated the performance of the proposed method by numerical experiments.

研究分野：統計科学

キーワード：圧縮センシング スパースモデリング 交差検証

1. 研究開始当初の背景

知りうることと知りえないこととを峻別することは重要である。いうまでもなくその重要性は、ビッグデータ時代と呼ばれる現代においても損なわれることはない。ビッグデータ時代が抱える最大の問題の一つは、データ量が膨大であるためにデータベース構築は固よりデータを逐一に保存することが難しくなっていることである。「知りうること」を知らないと、過剰なデータ取得により、この問題を助長することになる。一方で、「知りえないこと」を知らないと別の大きな問題が生じる。ビッグデータといえどもデータ量が不十分であることもある。このような場合、どんなに優れた情報処理手法を用いても所望の知識を得ることはできないだけでなく、時として誤った情報を供給してしまうことになり問題である。現有データのみを頼りに情報処理の成否を判断する枠組みを構築し、その限界を解明することが求められる。

自然科学をはじめとする実験・計測を要する分野では、効率的にデータを取得する方法が求められている。中でもスパースモデリングによる圧縮センシングは、分野を問わず注目を集めている。スパースモデリングとは対象を適切な基底で表現したときその係数のほとんどがゼロになり、非零係数が少数 (= スパース) であると仮定することである。スパースモデリングを用いると、実質的に推定すべき変数の数が少ないため、必要なデータの量をより少なくすることができる。このような考え方は圧縮センシングと呼ばれる。本研究開始までに、研究代表者が物質科学の代表的な計測手法である走査トンネル顕微・分光法に応用し、物質の表面状態の分散関係を明らかにする準粒子干渉実験を効率化するなど、スパースモデリングによる圧縮センシングの考え方がいろいろな実験・計測に広まりつつある。

2. 研究の目的

スパースモデリングに基づく圧縮センシングは実験・計測を効率化するための汎用性が高い考え方である。しかしながら、スパースモデリングを用いるといえども、データが少なすぎる場合に圧縮センシングは失敗する。ここで問題となるのは、どういう場合に圧縮センシングが成功していてどういう場合に圧縮センシングが失敗するのかということである。圧縮センシングの限界が曖昧であることにより、実験・計測者にとっては、手元のデータで十分なのか、さらなるデータ取得なのかという判断が難しい。本研究では圧縮センシングの限界を解明し、その限界を打破する方法を開発することを目的とする。

これまでの圧縮センシングの研究には、人工データを用いた数値実験や、実際に実データを用いるとしても標準試料を用いる場合が多かった。これらの場合に共通することは、人工データや標準試料など正解が良く知られていることが挙げられる。このような場合には、正解と比較することにより情報処理の成否を判断することができる。しかしながら、現実的な場合には、正解は知らないため、このような判断は難しい。

圧縮センシングの限界を解明する理論的な研究は皆無ではない。たとえば Donoho and Tanner は高次元幾何学を用いて、Candes and Tao は制限等長性を用いて、対象がスパースであれば、基底追跡法と呼ばれるアルゴリズムを用いて圧縮センシングが成功することを示した。基底追跡法は多項式時間で実行できるという簡便さもあり、これらの研究をきっかけにスパースモデリングによる圧縮センシングが広まった。しかしながら、この考え方をを用いて情報処理の成否を判断するとしても、対象が実際にスパースであることを別の方法で確認しなければならない。本研究では、対象やそのスパース性が未知である状況で圧縮センシングの成否を判断する方法を開発する。

3. 研究の方法

本研究では、統計力学の分野で発展してきた相転移の描像を活用することにより、圧縮センシングの成否を判断する方法を開発する。相転移とは、温度を下げると水が氷になるように、物質の状態が不連続的に変化することをいう。圧縮センシングについてもデータが増えると、あるデータ量で失敗から成功へと転移すると考えられる。実際、Kabashima et al. は、前述した Donoho and Tanner の数学的な議論を、相転移現象としてとらえ直し、圧縮センシングの適用範囲を広げることに貢献した。本研究の目的を達成するために、圧縮センシングの成否を分けるデータ量の転移点で起こる現象を明らかにし、手元のデータが成功相にあるか失敗相にあるか、すなわち十分なデータが得られているか、さらなるデータ取得が必要であるかを判断する方法を提案する。

本研究では、圧縮センシングにおける相転移現象を調べるために、統計科学の分野で伝統的に用いられてきた交差検証に着目した。交差検証とは、手元のデータをあえて訓練用と検証用との二つに分け、訓練データのみを用いて推定した結果により検証データを説明するときの誤差を調べることにより、推定の良し悪しを評価する方法である。交差検証は、あらかじめ正解を必要

とすることなく実行できることが利点である。また、スパースモデリングに限らずあらゆる推定手法と組み合わせて用いることができるので、本研究で開発する手法が普遍的な方法論となることが期待される。

4 . 研究成果

本研究で得られた主な成果は、ノイズ無し計測により得られるデータに対して、基底追跡法を用いて解析を行うという、一般的な圧縮センシングのシナリオにおいて、交差検証を活用することにより圧縮センシングの成否を診断する手法を開発したことである。本研究の特徴的な点は、交差検証において手元のデータを訓練データと検証データとに分けるときの、その比率を変えた時の交差検証誤差の振る舞いを調べたことである。統計力学の分野で発展したレプリカ法を用いて交差検証誤差の熱力学的極限における振る舞いを調べた。その結果、圧縮センシングの成否を分ける転移点上に手元のデータがあるとき、訓練データと検証データとの比率に対して交差検証誤差が冪的に振る舞うことを示した。また、データ量が転移点より大きいときと小さいときとで定性的に異なる振る舞いをすることを示した。つまり、データ量が大きくて圧縮センシングが成功しているときは交差検証が指数的な振る舞いをするのに対し、データ量が小さくて圧縮センシングが失敗しているときは、検証データが少ない極限において交差検証誤差が有限な値として残ることを明らかにした。この交差検証誤差の定性的な振る舞いの差を利用することにより、正解を知らないという現実的な状況で圧縮センシングの成否を判断する手法を提案した。提案した枠組みの性能を数値実験により示すとともに、基底追跡法の他のスパースモデリングのアルゴリズムに対しても同じ交差検証の枠組みを用いて圧縮センシングの成否を判断できることを数値実験により示した。

5. 主な発表論文等

〔雑誌論文〕 計12件（うち査読付論文 11件 / うち国際共著 0件 / うちオープンアクセス 8件）

1. 著者名 中西(大野)義典、福島孝治	4. 巻 54
2. 論文標題 走査トンネル分光法の圧縮センシング 計測の効率に関する限界と可能性	5. 発行年 2019年
3. 雑誌名 固体物理	6. 最初と最後の頁 343-351
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 無
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -
1. 著者名 Nakanishi-Ohno Y, Hukushima K	4. 巻 1036
2. 論文標題 Data-driven diagnosis for compressed sensing algorithms	5. 発行年 2018年
3. 雑誌名 Journal of Physics: Conference Series	6. 最初と最後の頁 012014 - 012014
掲載論文のDOI (デジタルオブジェクト識別子) 10.1088/1742-6596/1036/1/012014	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -
1. 著者名 Obuchi Tomoyuki, Nakanishi-Ohno Yoshinori, Okada Masato, Kabashima Yoshiyuki	4. 巻 2018
2. 論文標題 Statistical mechanical analysis of sparse linear regression as a variable selection problem	5. 発行年 2018年
3. 雑誌名 Journal of Statistical Mechanics: Theory and Experiment	6. 最初と最後の頁 103401
掲載論文のDOI (デジタルオブジェクト識別子) 10.1088/1742-5468/aae02c	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -
1. 著者名 Nakanishi-Ohno Yoshinori, Hukushima Koji	4. 巻 98
2. 論文標題 Data-driven diagnosis for compressed sensing with cross validation	5. 発行年 2018年
3. 雑誌名 Physical Review E	6. 最初と最後の頁 52120
掲載論文のDOI (デジタルオブジェクト識別子) 10.1103/PhysRevE.98.052120	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

〔学会発表〕 計27件（うち招待講演 11件 / うち国際学会 3件）

1. 発表者名 中西（大野）義典
2. 発表標題 LASSO変数選択による予測と発見のデータ駆動診断
3. 学会等名 日本物理学会2018年秋季大会
4. 発表年 2018年

1. 発表者名 Nakanishi-Ohno Yoshinori
2. 発表標題 Data-driven diagnosis for compressed sensing algorithms
3. 学会等名 International Meeting on "High-Dimensional Data-Driven Science"（国際学会）
4. 発表年 2017年

1. 発表者名 中西（大野）義典
2. 発表標題 交差検証による圧縮センシング成否診断法の開発
3. 学会等名 日本物理学会2017年秋季大会
4. 発表年 2017年

1. 発表者名 Nakanishi-Ohno Yoshinori
2. 発表標題 Data-driven diagnosis for compressed sensing with sparse modeling
3. 学会等名 Workshop 'Development of next-generation quantum material research platform'（招待講演）（国際学会）
4. 発表年 2017年

1. 発表者名 中西(大野)義典
2. 発表標題 LASSOに関する交差検証誤差の訓練 / 検証データ量比依存性
3. 学会等名 日本物理学会第73回年次大会
4. 発表年 2018年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
---------------------------	-----------------------	----

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------